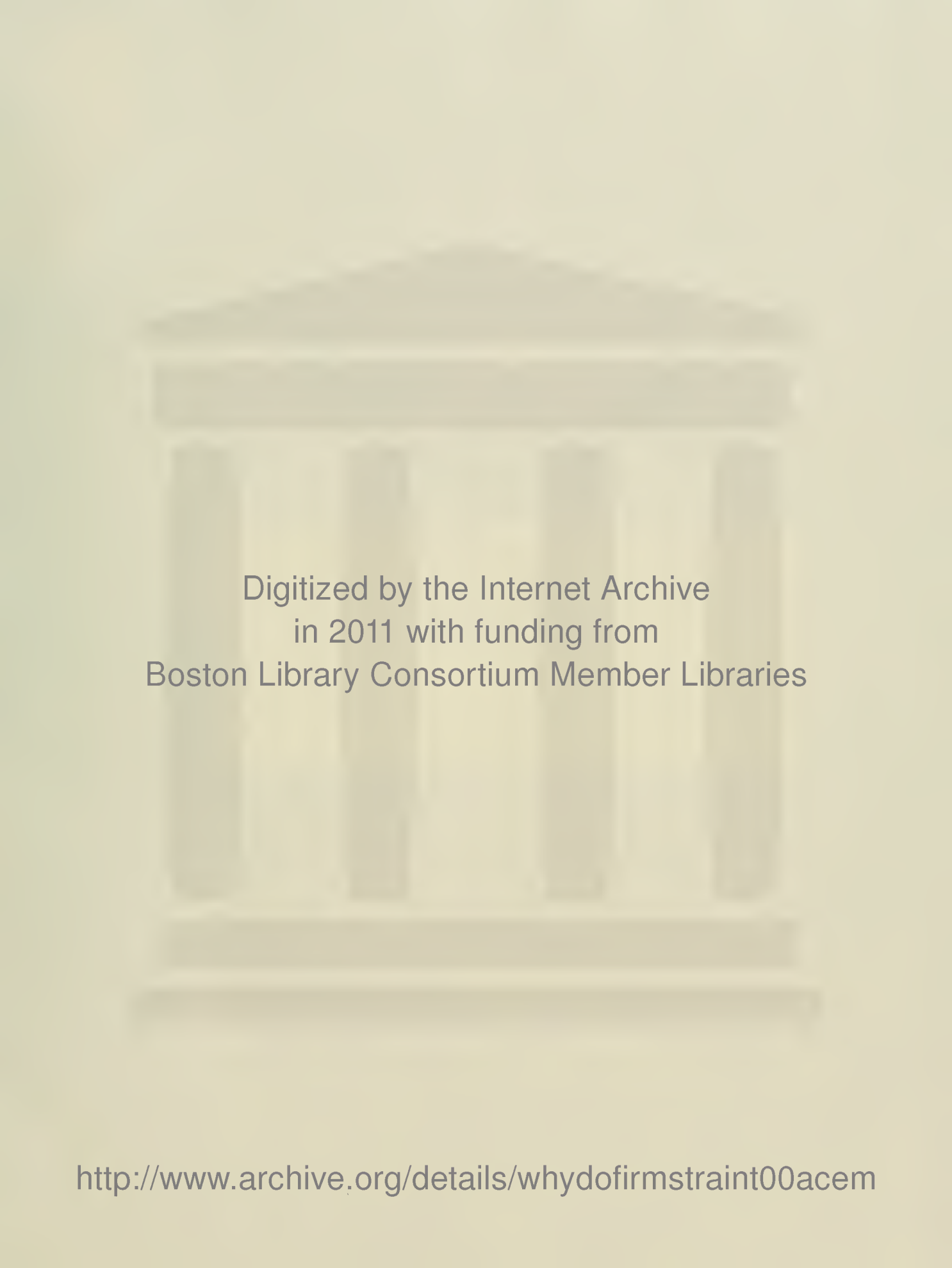
MASSACHUSETTS
INSTITUTE of TECHNOLOGY
LIBRARIES


Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/whydofirmstraint00acem

DEWEY

HB31
.M415
no.96-7

working paper
department
of economics

WHY DO FIRMS TRAIN? THEORY AND EVIDENCE

Daron Acemoglu
Jorn-Steffen Pischki

96-7 Feb. 1996

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

WHY DO FIRMS TRAIN? THEORY AND EVIDENCE

Daron Acemoglu
Jorn-Steffen Pischki

96-7 Feb. 1996

MASSACHUSETTS INSTITUTE

MAR 13 1996

LIBRARIES

Why Do Firms Train? Theory and Evidence

Daron Acemoglu
MIT

Jörn-Steffen Pischke*
MIT

February, 1996

Keywords: Asymmetric Information, General Human Capital, German Apprenticeship System, Monopsony, Training.

JEL Classification: D82, J24, J42.

We thank Bob Gibbons, Kevin Lang and Lones Smith for useful comments. This research has been supported by a grant from the German Federal Ministry for Research and Technology to the World Economy Laboratory at MIT. The data used in this paper have been obtained from the German Zentralarchiv für Empirische Sozialforschung at the University of Köln (ZA). The data for the study "Qualifikation und Berufsverlauf" were collected by the Bundesinstitut für Berufsbildung and the Institut für Arbeitsmarkt- und Berufsforschung and documented by the ZA. Neither the producers of the data nor the ZA bear any responsibility for the analysis and interpretation of the data in this paper.

Why Do Firms Train? Theory and Evidence

Abstract

This paper offers and tests a theory of training whereby workers do not pay for general training they receive. The crucial ingredient in our model is that the current employer has superior information about the worker's ability relative to other firms. This informational advantage gives the employer an ex post monopsony power over the worker which encourages the firm to provide training. We show that the model can lead to multiple equilibria. In one equilibrium quits are endogenously high, and as a result employers have limited monopsony power and are willing to supply only little training, while in another equilibrium quits are low and training high. We also derive predictions from our model not shared by other explanations of firm sponsored training. Using microdata from Germany, we show that the predictions of the specific human capital model are rejected, while our model receives support from the data.

Keywords: Asymmetric Information, General Human Capital, German Apprenticeship System, Monopsony, Training.

JEL Classification: D82, J24, J42.

1. Introduction

A large portion of human capital accumulation in the form of training and on-the-job learning takes place inside firms. Becker (1965) and Mincer (1972) provide a systematic explanation of training investments and the associated wages of workers. This standard theory draws a crucial distinction between general and firm-specific training. General training will increase workers' wages in a range of employment opportunities, thus it is the worker who has to pay for general training. The firm should only pay for the firm-specific component of training which does not help the worker receive higher wages elsewhere. However, these predictions seem to be at odds with reality. In Germany, firms voluntarily offer apprenticeships to workers entering the labor market. Although general skills are an important component of these programs, much of the financial burden is borne by the firm (see Noll et al., 1983). The same is true in Japan where training and mentoring within companies is more important in many areas than formal schooling (e.g. there are no business schools in Japan). Many labor economists also believe that U.S. firms provide some general training to their workers while bearing some of the initial outlays, for example in the form of tuition assistance or as in-house training in general skills (see Osterman for a discussion of the U.S. evidence).

In order to explain firms' investments in general skills, some labor market imperfection must exist so that the mobility of workers is restricted and that employers earn rents on trained workers. The mechanism we propose in this paper is the adverse selection story also analyzed by Waldman (1984), Greenwald (1986) and Gibbons and Katz (1991) among others. The talents and skills of young workers who arrive in the market are in general unknown. Some will be more productive and able at tasks in which they choose to specialize than others. The early years of the worker's career constitute the period in which more information is accumulated about his ability. However, most of this information is gathered by the employer and not transmitted to outsiders. The presence of this asymmetric information has already been noted in the literature. The crucial step we take is to note the link between the firm's ex post *monopsony power* over the worker arising from this informational advantage and training incentives.

Given that the firm is able to obtain part of the marginal product of the worker, it also has an interest in increasing this marginal product by investing

in the worker's human capital. Workers, on the other hand, may not be willing to pay for general skills themselves because they realize that at least part of the returns will be appropriated by the firm. As a result, in stark contrast to the standard Beckerian model of training, a firm which wants to attract workers to its apprenticeship program may have to pay not only for the training but also offer apprentices a training wage above their initial productivity. There is no need to resort to credit constraints of workers for this to be true.

The failure of the U.S. economy to generate as much training as Germany or Japan is sometimes blamed on the higher turnover in the U.S. (e.g. Blinder and Krueger, 1991)[1]. However, if training is general to a large extent, this statement cannot be reconciled with the standard model of training: with or without quits, firms will have to pay the full marginal product of the worker in the outside markets. Thus, exogenous differences in turnover alone do not explain the different practices in financing general training. In our model quit rates are endogenous and play an important role for equilibrium wages and training levels. Because quit rates determine the composition of the second-hand market for workers, they influence the wages that outside firms are willing to pay and thus the degree of monopsony that the current employer has over the workers. This channel will also imply that our model is able to generate multiple equilibria, or more importantly, very different equilibrium outcomes resulting from small differences in underlying parameter values. For example, an economy could be characterized by high training and low quits. Because workers who quit receive low wages, the apprenticeship firm will have substantial monopsony power which supports a high level of training. In a different equilibrium the same economy will have low training and high quits. In this equilibrium, quitters receive relatively high wages, leave their firms with a high frequency, and the firm will only enjoy moderate monopsony power. Paradoxically, the equilibrium with high quits, which achieves a better allocation of workers to jobs, may be less efficient since the level of training is lower.

A model of training would be somehow incomplete without an explicit treatment of the possibility of active poaching or raids by other firms, especially given the prevalence of this type of behavior in the U.S. labor market. However, here too

[1]Turnover is certainly much higher in the U.S. than in Germany. Topel and Ward (1992) report that in the U.S. labor market, the median number of jobs that a worker with 10 years of experience holds until the end of his career is 6. We calculate the same statistic to be 2 in our 1979 data and 1 in our 1985 data for Germany. However, it is important to emphasize that in our model, the turnover rate is not exogenous, instead turnover, training and wages will be simultaneously determined in equilibrium.

asymmetric information is more powerful than previously envisioned. We show that raids lead to an extreme form of the winner's curse which often makes this strategy unprofitable, and as a result the monopsony power of the apprenticeship firm is preserved. Nevertheless, in the presence of raids there will be a unique equilibrium with no more training than in the case without poaching.

While our theory, at least to us, has obvious appeal, a real test involves deriving new and different predictions and confronting them with data. We do this in the second part of the paper. In particular, we look at the wages of German apprentices who stay in their firm, who are laid off or quit their firm voluntarily, and compare them with the wages of apprentices who quit for *exogenous* reasons, that is to go to the military. West Germany has had a draft system since 1957, but not all males get drafted. There is some screening on physical fitness, but otherwise the selection into the military is random. Our model predicts that military quitters should have higher wages than voluntary quitters (and laid off workers) which is confirmed in the data. This test has an obvious similarity to Gibbons and Katz's (1991) empirical investigation of adverse selection in the U.S. labor market. However, we also draw a stronger prediction from our model that is not shared by other models. Because the firm has monopsony power over workers who stay, they are paid below their marginal product. In contrast, military quitters, thanks to their exogenous reason of separation, are freed from this monopsony power. While they will be *less able* than stayers on average, they can have *higher wages*. In fact, we find that their wages are very close to or above those of stayers. Further, using proxies for ability available in our data set, we also present evidence which suggests that, despite their higher or equal earnings, military quitters are *less able* than stayers! We also contrast our empirical findings to the implications of a standard human capital model where there are no information asymmetries and training is firm-specific, and we show that they cannot easily be reconciled with this alternative.

The remainder of the paper is organized as follows. In the next section we lay out the basic model. Section 3 discusses extensions to the model, in particular the potential of raids on trained workers by outside firms. We also compare our model to related literatures. Section 4 outlines the testable implications of the model. Section 5 describes the data and section 6 discusses our empirical findings. In section 7 we conclude by drawing comparative lessons for the U.S.

2. Theory: The Basic Model

We start with the simplest model to obtain the sharpest predictions. The next section deals with extensions aimed at demonstrating the robustness of the model.

2.1. Description of the Environment

The world lasts for two periods. Firms can hire workers at the beginning of either period. At all points, firms have constant returns to scale. All firms and workers are risk-neutral. Also, there is no discounting between periods. During the first period, firms can offer training to the workers they hire. The amount of this training is denoted by t. With or without training workers produce nothing during the first period.

In the second period each worker produces

$$y = \alpha(t)\eta \tag{2.1}$$

where $\alpha(t)$ is the general human capital and η is the ability of the worker. The assumption that all human capital is general is of course extreme, but it will demonstrate that none of our results are driven by firm-specific human capital. It is important for our results that ability and training are complementary; the multiplicative specification of (2.1) captures this in a simple fashion. Regarding $\alpha(t)$ we make the following assumption:

Assumption 1 $\alpha(t)$ is twice continuously differentiable, strictly increasing and concave with $\lim_{t\to 0}\alpha'(t) = \infty$ and $\lim_{t\to \bar{t}}\alpha'(t) = 0$ for some $0 < \bar{t} < \infty$.

The cost of training each worker is given by t. Ability is distributed according to a continuous and differentiable distribution function $F(\eta)$ with support $[0, \eta^{\sup}]$. In the first period, no one knows the ability of a worker, though the distribution from which these abilities are drawn is common knowledge. At the end of the first period, the firm learns the ability of each worker it employs. Note quite importantly that this is independent of whether and how much training is offered to the worker. *Training neither facilitates nor leads to learning.* The reason firms train will *not* be because training enables them to learn better.

The sequence of events is as follows. In the first period, firms decide how many workers to hire and how much training to give them. Throughout the analysis, we will assume that each firm employs a continuum of workers. At this stage the firm does not observe the ability of the workers. At the end of period 1, the firm

observes the ability of all its employees, and it makes a unique wage offer to these workers, which we denote by $w(t)$. At this point, it may also desire to lay off some of these workers.[2] However, the firm is aware that some of the workers it wants to keep will nevertheless leave. The reason for quits is because the workers realize that they are not getting on with their colleagues or supervisor, or because their spouse moves to a different city or simply because they have realized that they are more suited to another profession or environment. To capture these and other reasons for voluntary quits we assume that with probability λ each worker receives a *negative* taste shock θ. The disutility shock θ has a distribution function $G(\theta)$. This shock is never observed by any agent other than the worker in question.

We now make some assumptions about the distribution functions:

Assumption 2 $G(x)$ is twice differentiable and is log-concave in x.

Assumption 3 $F(x)$ is twice differentiable and $\int_x \eta dF(\eta)$ is log-concave in x.

Log-concavity ensures well-behaved second-order conditions and comparative static results (Burdett, 1979) and has become a standard assumption in models of this type. It simply requires that the log of the function in question is concave, most well-known distribution functions including uniform and normal satisfy this requirement. These assumptions are by no means necessary for our results, but they simplify the exposition and proofs considerably.

Workers who quit receive wage offers from outside firms. These firms do not observe the ability level of the worker prior to making a binding wage offer, thus offers are based on their beliefs about the workers. Another important factor is that outside firms never observe whether a worker has quit voluntarily or has been laid off. However, the exact training level that the worker has received is always observable.[3] We denote the outside wage of a worker who quits or is laid off and has a level of training t by $v(t)$.

[2]Throughout the paper we assume that the output of the workers is not verifiable (for instance, only the overall output of the firm is observable to outside parties), thus neither the first employer nor future employers can write contingent wage contracts.

[3]This is a reasonable assumption since outside firms know how much training different apprenticeship firms offer but can never ascertain whether the worker has really quit or anticipating that he would be laid off, he jumped the gun. For instance, consider the case where quits and layoffs were distinguishable and the worker learned his type an instant before the firm, he would immediately quit and thus there would be no layoffs, hence distinguishing quits and layoffs would convey no useful information. A theory of training whereby outside firms do not observe the training decisions of the apprenticeship firm can be developed in a very straighforward manner (e.g. Chang and Wang, 1994). However, as well as being less realistic, especially for the

Finally, we impose a free-entry condition on firms at all points in time. That is, in equilibrium no firm will earn positive profits. As we will see this may imply that apprentices (workers in period 1) need to be paid a positive wage even though they are not productive during this period. We denote this wage by W. An assumption we make is that firms only compete for workers by offering a training wage W, which is not contingent on the level of training, and then unilaterally decide the level of training to maximize profits. This basically corresponds to assuming that contracts on training levels are hard to enforce, for instance, the amount of training is not perfectly verifiable. An alternative assumption is that firms offer a package consisting of training levels and wages $[W, t]$ to attract young workers. The model is solved under this assumption in section 3.

2.2. The Competitive Equilibrium (No Information Asymmetries)

The model we have is non-competitive due to the informational asymmetries. If the ability level η of each worker is observed in the second period by all firms, then we obtain Becker's standard model.[4] In this case, outside wages of workers who quit and have ability η and training t are given by $v(t) = \alpha(t)\eta$. Thus no firm will be able to retain its trained workers unless it pays this wage. As a consequence, the training firm enjoys no rents, and there will be no training unless the worker is willing and able to pay for his training. If we impose the credit constraint that $W \geq 0$, the model delivers $t^0 = 0$. On the other hand, if W is unconstrained, then we have the optimal amount of training that is

$$t^c = \arg\max \int \eta\alpha(t)dF(\eta) - t. \tag{2.2}$$

and the corresponding training wage is $W = -t^c$. We will contrast our solution with this competitive (Beckerian) equilibrium later.

2.3. Equilibrium With Adverse Selection

In the rest of the paper we will analyze the case in which $F(\eta)$ is non-degenerate and η is not observed by firms other than the current employer. We begin with

German case, such a model is driven by the same features as a model where human capital is firm specific, and therefore no real contrast to the standard Beckerian story.

[4]Obviously, the Beckerian model is also obtained in the limiting case where $F(\eta)$ is degenerate at some level $\tilde{\eta}$.

the assumption that there is no possibility of poaching or raids. We will return to this issue in the following section.

2.3.1. The Equilibrium Concept

The equilibrium concept we will use is Perfect Bayesian. We have at hand a large and complex game. The number of players is infinite, and there is a continuum of information sets that can be reached in the second period of the economy (depending on the training level of workers).[5]

An allocation is an equilibrium iff

1. All workers are taking the optimal quit decisions given the wage offer and the outside wage.

2. All apprenticeship firms (firms with trainees) are choosing the optimal lay-off, wage, and training rules, $\hat{\eta}^*(t), w^*(t), t^*$ and the optimal training wage W^*, taking as given the outside wage offers and the workers' optimal quit decisions at all points in the game (including the optimal layoff decisions for the training levels that do not arise in equilibrium).

3. All outside firms are making optimal offers, $v(t)$, to workers looking for a job in the second period given their beliefs regarding the ability of these workers.

[5]We have a game in which the informed players (apprenticeship firms) move before the uninformed players (outside firms) and in this type of games the concept of Perfect Bayesian Equilibrium strengthened by some signalling requirement is often used. However, it turns out that in our case the Perfect Bayesian concept is sufficient because we can determine the relevant beliefs in all information sets. An alternative way of expressing this is that all of the Perfect Bayesian Equilibria of this game satisfy the Intuitive Criterion of Cho and Kreps (1987) and even stronger refinements. The key to this "good performance" of Perfect Bayesian Equilibrium is that before the crucial information sets of the uninformed players are reached, the informed player has a decision that pins down the subsequent beliefs. More explicitly, the game-theoretic problem could be that the information set, in which workers with other levels of training ($t \neq t^*$) look for a job in the second period, is not reached. Thus it may be conjectured that $v(t)$ is free except at $v(t^*)$ which will then introduce a number of unreasonable equilibria. However, suppose that a firm deviates and chooses a training level $t \neq t^*$, are beliefs free? The answer is no because after this choice of training the firm has a choice of which workers to lay-off and here it will always find it profitable to layoff workers that have $\eta < w(t)/\alpha(t)$, thus beliefs of outsiders even when $t \neq t^*$ are determined.

4. The beliefs of outside firms are derived by Bayes' Rule in all information sets that are reached.

The set of functions and numbers $\{v(t), w^*(t), \hat{\eta}^*(t), t^*, W^*\}$ is an equilibrium if conditions 1.-4. above hold.

2.3.2. Analysis and Characterization of Equilibria

Denote the probability that a worker with training level t quits by $q[w(t), v(t)]$. A worker with training t who quits receives $v(t)$ whereas if he stays he will get $w(t)$, but he will also suffer disutility θ with probability λ. The optimal decision is to quit iff

$$v(t) - w(t) + \theta \geq 0.$$

Therefore, the probability of quitting is

$$q[w(t), v(t)] = \lambda[1 - G[w(t) - v(t)]].$$

Now recall that the firm is forced to pay the same wage to all remaining workers of the same training level. Hence the firm's layoff decision will take a simple cut-off form whereby all workers above a skill level $\hat{\eta}(t)$ will be kept.

Since the firm offers W first and then decides how much training to offer and chooses the retention and wage policies, the problem can then be written as

$$\max_{w(t),\hat{\eta}(t),t} \Pi = [1 - q[w(t), v(t)]] \int_{\frac{w(t)}{\alpha(t)}}^{\infty} [\alpha(t)\eta - w(t)] dF(\eta) - t - W. \tag{2.3}$$

In words, the firm takes the wage that other firms offer to quitters, $v(t)$, as given and maximizes its profit by choosing the cut-off function for layoffs $\hat{\eta}^*(t)$, wage function to be paid to all workers $w^*(t)$, and the training level t^*. In equilibrium, we only observe $\hat{\eta}^*(t^*)$ and $w^*(t^*)$. Nevertheless, the rest of these functions need to be determined as part of the equilibrium strategies in order to pin down the off-the-equilibrium-path behavior. We will first deal with the determination of $\hat{\eta}^*(t)$ and $w^*(t)$ and then return to t^*.

The first order conditions are (second order conditions are satisfied given assumptions 2 and 3, see proof of Proposition 1):

$$\hat{\eta}^*(t) = \frac{w^*(t)}{\alpha(t)}, \tag{2.4}$$

$$-[1-q(w^*(t),v(t))][1-F(\hat{\eta}^*(t))]-\frac{\partial q[w^*(t),v(t)]}{\partial w(t)}\int_{\hat{\eta}^*(t)}^{\infty}[\alpha(t)\eta-w^*(t)]dF(\eta)=0, \tag{2.5}$$

for all t. The optimal level of t, t^* is given by:

$$[1-q[w^*(t^*),v(t^*)]]\int_{\hat{\eta}^*(t^*)}^{\infty}\alpha'(t^*)\eta dF(\eta)-1=0. \tag{2.6}$$

Next, in order to attract workers, the firm will be willing to offer a training wage W as high as the maximized value of (2.3). Thus

$$W^*=\max\{0;[1-q[w^*(t^*),v(t^*)]]\int_{\hat{\eta}^*(t^*)}^{\infty}[\alpha(t^*)\eta-w^*(t^*)]dF(\eta)-t^*\}. \tag{2.7}$$

since these wages cannot be negative (by the credit constraint of the workers). Although our model works equally well without credit constraints, to impose the restriction $W \geq 0$ is a good way of highlighting our differences from the competitive (no information asymmetries) model. Proposition 3 below will show that in equilibrium W will sometimes be strictly positive, thus this restriction will not bind.

Finally, we need to determine how outside wages for quitters are determined. Due to free-entry these workers will be paid their expected marginal product, which by Bayes' rule gives:

$$v(t)=\frac{q[w^*(t),v(t)]\alpha(t)\bar{\eta}+[1-q[w^*(t),v(t)]]\int_0^{\hat{\eta}^*(t)}\alpha(t)\eta dF(\eta)}{q[w^*(t),v(t)]+[1-q[w^*(t),v(t)]]F(\hat{\eta}^*(t))} \tag{2.8}$$

Equations (2.4), (2.5), (2.6), (2.7) and (2.8) determine $v(t), w^*(t), \eta^*(t)$, t^* and W^*.

Proposition 1. *(i) An equilibrium, with $w^*(t)$, $v(t)$ and $\hat{\eta}^*(t)$ continuously differentiable, exists and is characterized by (2.4)-(2.8).*

(ii) Denote the set of equilibrium training levels by T^{np}. Then $\forall\, t^a \in T^{np}$, $t^0 = 0 < t^a < t^c$ where t^0 and t^c are the training levels in the economy without informational asymmetries and, respectively, with and without credit constraints.

Proof: See the appendix.

Part (i) establishes the existence of a well-behaved equilibrium. For the purposes of our paper, the more important result is in part (ii) which characterizes some important properties of the equilibrium, namely that the amount of training is always positive but less than first-best.

Because in our economy the firm will have an ex post monopsony power over its employees, it will be able to capture a portion of their marginal product. Given the multiplicative structure of ability and training in equation (2.1), this also implies that the firm can increase the monopsony rents by training. Therefore, the model always generates a positive amount of training. Clearly, this is higher than the training level of zero (t^0) the competitive model would achieve when workers are unable to pay for their own training. As commented above, the multiplicative structure of ability and training per se is not crucial for this result. Rather the important feature is that $\alpha(t)$ and η are complements. In contrast, if (2.1) were additive, i.e. $y(t) = \alpha(t) + \eta$, then the firm would still have ex post monopsony power, but would only obtain the proportion of the worker's marginal product coming from η. All increases coming from $\alpha(t)$ would be reflected in the worker's outside wage.

Remember that the firm lays off some workers and other workers quit. Training of these workers is costly but the firm only reaps the benefits from workers who stay with the firm. Therefore, in this non-competitive structure, due to the leakage arising from layoffs and quits, the training level is also less than the unconstrained, full-information first-best (t^c).

Let us turn to the characterization of equilibrium next.

Proposition 2. *(i) There can exist multiple equilibria.*

(ii) Suppose there exist multiple equilibria and that Assumption 2 holds, then for any two equilibria a and b, we have that $v^a(t) > v^b(t) \Leftrightarrow q^a[w^a(t), v(t)] > q^b[w^b(t), v(t)] \Leftrightarrow t^a < t^b$.

Proof: See the appendix.

Our model therefore endogenously generates two equilibria for the same parameter values; one with high training and low turnover and the other with low training and high turnover.

Let us explain the intuition of this result and then give an example. The RHS of (2.8) is increasing in the probability of quitting q. That is, at a given level of

training t, $\partial v(t)/\partial q > 0$. More voluntary quits imply that the quality of the labor pool in the secondary market is higher. To see this consider the case in which $q = 0$. Then all workers who are looking for a job must have been laid off, and thus they must have a level of ability less than $\hat{\eta}$. In contrast, when $q > 0$, some of the quitters may be of very high ability but leave voluntarily because they were not happy with their jobs. Hence, the higher the quit probability, the higher the outside wage. Furthermore, from the definition of q we see that[6]

$$\frac{\partial q[w^*(t), v(t)]}{\partial v(t)} = \lambda g[w^*(t) - v(t)] > 0.$$

Higher outside wages imply that workers who are unhappy can leave more easily, and as a result voluntary quits are higher.

Therefore, at a given level of training, the equilibrium wages and quit behavior can be determined as the intersection of two upward sloping curves, hence the multiplicity. Also, with higher quits profits due to training are lower because the firm has a smaller ex post monopsony power over trained workers, and this leads to less training in equilibrium.

Proposition 3. *There exists an open set of parameter values such that W^* is strictly positive.*

Proof: See the Appendix.

This is perhaps the most interesting result of our paper. It emphasizes how our model differs from the standard theory outlined in section 2.2. In the standard theory, workers have to pay for their general training. Here, they get paid an additional amount. Thus even if workers are not credit constrained, *they will not pay for their training.* The underlying reason is naturally the firm's ex-post *monopsony* power over trained workers. Depending on the equilibrium beliefs, post-training wages may be so low that workers require compensation ex ante, that is $W^* > 0$.

Finally, a brief discussion of the efficiency issues is necessary. Since $\forall t^a \in T^{np}$, $t^a < t^c$, the training level is inefficiently low. This is due to the fact that the firm cares only about its own profits when it chooses the level of training (see (2.3)), whereas higher training also increases the return to the workers, especially to those who are laid off or choose to quit in the second period. When we allow

[6]The same is true for dq/dv, see the Appendix.

firms to attract workers by offering a combination of training wages and training levels in section 3, part, but not all, of this inefficiency will be avoided. Second, it is natural to ask whether an equilibrium with higher level of training Pareto dominates one with lower training. The answer is ambiguous (see the example below). To see the intuition note that since $t < t^c$, increasing the level of training is in general beneficial. However, an equilibrium with higher training also has lower quits. Therefore more workers who have received negative utility shocks and who would be better off finding a new job are forced to stay in their old firm. This means that the allocation of workers to jobs is more distorted. In general it is not possible to ascertain which of these two forces will dominate.

2.3.3. An Example of Multiplicity

In this subsection we give a parameterized example of multiple equilibria. We first show how multiple wages and quit levels can arise as part of an equilibrium at a given training level, and then show how this leads to multiple equilibrium training levels.

Consider the following specific features. The distribution of the taste shock $G(.)$ is uniform over $[0,1]$. Training takes only two values, $t \in \{0, C\}$, where $C > 0$ is the cost of training. Also $\alpha(0) = 1$ and $\alpha(C) = \alpha > 1$. Ability also has only two values: with probability p, $\eta = 0$ and with probability $1-p$, $\eta = 1$.[7]

Suppose a firm decides to train. Then, profits with training $t = C$ are given by

$$\pi(t = C) = (1-p)(w-v)(\alpha - w) - C \tag{2.9}$$

Thus, all workers with $\eta = 0$ are laid off and the optimal wage offer to remaining workers maximizes equation (2.9). This yields $w = (\alpha + v)/2$. Wages for quitters and laid off workers are given by

$$v = \frac{(1-p)(1-w+v)\alpha}{p + (1-p)(1-w+v)} \tag{2.10}$$

Combining the expression for the wage w and (2.10) we get

$$(1-p)v^2 - 2[(1-p)\alpha - 1]v + (1-p)\alpha(\alpha - 2) = 0$$

[7]Note that in our main analysis we assumed $F(\eta)$ to be continuously differentiable. This was adopted for simplicity and our results hold with $F(\eta)$ having atoms as is the case in this example.

which has two positive real roots as long as $\alpha > 2$ and $(1-p)p\alpha < 1/2$. Thus for the same level of training of the workers, there can be two equilibrium levels of wages w and v and quit behaviors. Furthermore, with lower wages for the quitters, quits are lower and profits are higher and depending on the value of training cost, C, we can also have multiple training equilibria.

Consider for instance $p = 0.2$, $\alpha = 3$. If the firm trains, we get two wage equilibria with

1. $v = 2.00$, $w = 2.50$, $q = 0.50$, $\pi(t = C) = 0.20 - C$
2. $v = 1.50$, $w = 2.25$, $q = 0.25$, $\pi(t = C) = 0.45 - C$

Furthermore, we can easily incorporate into this example a small group of workers quitting their apprenticeship firms for exogenous and *publicly observable* reasons (see section 4 for the relation of this to our empirical test). The wages of these workers in both scenarios will be $v_m = 2.4$, the mean productivity of trained workers. Thus, in both cases these workers earn more than voluntary quitters and laid off workers, and in one of the equilibria they earn more and in the other less than the stayers. However, on average they are always *less* able than stayers!

To develop the example further, if the training cost $C = 0.3$, then we also have two equilibria with different training levels. In the first, scenario 1 is expected and no workers are trained. In the second, scenario 2 is anticipated after the apprenticeship, and firms knowing that they will be able to recoup all the cost of training in the second-period, offer training to all the workers:

1. $v = 0.78$, $w = 0.89$, $q = 0.89$, $t = 0, W = 0.01$
2. $v = 1.50$, $w = 2.25$, $q = 0.25$, $t = C, W = 0.15$

In both equilibria the "training wage" is positive, but it is actually higher when training is positive . That is, in stark contrast to the Beckerian conventional wisdom that predicts that workers should pay for their general training, here workers are being paid an additional amount to take part in an apprenticeship. This is because of the informational advantage of the incumbent firm which leads to a low wage in period 2. Free entry ensures that the firm has to pay for this profit opportunity at date 1.

Finally, it is easy to verify that with the parameter values we have chosen all workers (before they know their η and θ) prefer to be in a high training/low quit equilibrium rather than a low training/high quit equilibrium. However, if

we change parameter values such that $C = 0.1$, then with or without high quits, training will be an equilibrium. In this case, workers prefer the high quit equilibrium with less monopsony power for the firm since in this equilibrium they will not be forced to stay in jobs to which they are not well-suited.

3. Extensions and Discussion

3.1. Raids on Trained Workers

So far we have only allowed outside firms to hire quitters which can be thought of as *passive poaching*. The alternative is of course *active poaching* or *raids* whereby a firm makes an offer to a currently employed worker. To analyze these possibilities, we now add more structure to our model. As before, outside firms observe the training level of workers employed by another firm, but observe neither this worker's ability η, nor his willingness to quit the current job, θ. Again θ in not observed by the current employer either. A raider makes an offer $v_p(t)$ to all workers with training level t at a certain apprenticeship firm. This offer specifies the wage that the worker will receive if employed by the raider regardless of his ability. We assume that this offer is observed by the apprenticeship firm who can then make a counter offer which we denote by $w_R(t)$. At this point we truncate the bidding war and the worker has to decide which of these offers to take.[8]

To obtain clean results, we distinguish two separate scenarios.

3.1.1. Scenario 1: Poaching offers at the end of period 1.

In this case, raiders have to make the offer without knowing which workers will be laid off by the training firm. They will also expect the apprenticeship firm to make a counter-offer. This will lead to an extreme form of *winner's curse*. In the case where $\lambda = 0$, the raider will only obtain a worker when its bid is above the productivity of the worker. Using this idea, we can establish the following:

Lemma 1. *If the poaching offer is made at the end of period 1, then* $v_p(t) > \max V^{np}$ *is not profitable where* V^{np} *is the set of equilibrium outside wage functions without poaching.*

[8]We conjecture that an ascending (English) auction between the outside firm and the apprenticeship firm will not change the results in a substantial way, athough we have not formally proved this.

With the poaching offer at the end of period 1, the firm will get all workers who would be otherwise laid off and some of the high productivity workers with $\theta > 0$. However, these are precisely the workers who would have chosen to quit had the outside wage been $v_p(t)$. But then (2.8) determines which outside wage functions break even. Thus, the raider never wants to make an offer above the maximum *equilibrium* outside wage offer, i.e. an offer that violates (2.8).

As a consequence, while the possibility of raids does not make training impossible, it still discourages it. This result arises because raids destroy the multiplicity of equilibria; in particular the equilibrium with low quits does not survive the introduction of raids. To see this, suppose that without raids there were multiple equilibria such that a worker with training t' could receive either $v(t')$ or $\hat{v}(t') > v(t')$. What maintains both of these outside wages as equilibria is that at the lower wage, $v(t')$, few workers quit because they expect low wages and this lowers the average quality of the workers in the second-period labor market. Yet, if all these workers quit, the outside wage would be high, $\hat{v}(t')$, and quitting would be preferred. This implies that by offering $\hat{v}(t')$ the raider could induce all these workers to quit, thus the equilibrium with lower training and higher outside wages would always be selected.

As a consequence, it follows that:

Proposition 4. *With poaching offers at the end of period 1, there exists a unique equilibrium level of training. Let the set of equilibrium training levels without poaching be T^{np}, and let the set of outside wage functions be V^{np}. Then the equilibrium training level with poaching is $t^p = \min T^{np}$ and the equilibrium outside wage offer is $v^p(t) = \max V^{np}$.*

Proof: See the appendix.

3.1.2. Scenario 2: Poaching offers can be made at any time.

In this case the raider firm can wait for layoffs and then make a poaching offer to the remaining trained workers. In the previous case the poaching offers were "timid" since the raider was afraid of attracting the workers who would otherwise be laid off. Now this is no longer true and raiders are willing to offer higher wages. The game is now much more complicated and a full analysis is beyond the scope of this paper. Here we can establish that our previous results do not change in fundamental ways.

The crucial assumption we make is that the number of workers each training firm lays off is observable. Because the workers who are not laid off will be of higher quality, quitters will want to wait for the more favorable poaching offer rather than be pooled with the laid off workers. This implies that there is a strategic element in the layoff decision of the training firm. By laying off fewer workers, the training firm can induce raiders to make lower offers. To see this consider the case where the training firm lays off half of its workers, then outside firms will deduce that the remaining workers are all above average ability. On the other hand, if the firm does not lay off many workers, then raiding firms will recognize that they will also attract more of the lemons.

To formalize this intuition consider the following sequence of events. The apprenticeship firm trains all workers, then offers a unique wage to all workers and lays off some. These actions are perfectly observed by outsiders. Then outsiders make raiding offers. The firm can respond to these raiding offers if it wishes. Finally, the workers decide which of these offers to take.

In this case, an equilibrium will take the following form. The training firm offers a zero wage to all of its workers and lays off none. This strategy has no cost for the firm since it knows that in equilibrium there will definitely be raids and thus it can get rid of the workers it does not want by not making counter-offers.[9] Then equilibrium is characterized by a raiding offer $v_p(t)$ such that all raiding firms make this offer and make zero profits. Then the firm optimally responds to $v_p(t)$ by offering $w_R(t) = \max\{v_p(t), w^*(t)\}$ to all workers that have $\eta \geq w_R(t)/\alpha(t)$ where $w^*(t)$ is as defined in section 2. To see why the apprenticeship firm's best response takes this simple form, note that if $v_p(t) < w^*(t)$, then the maximization problem of section 2 applies to the firm. In contrast, if $v_p(t) > w^*(t)$ and the incumbent firm wants to keep the worker in question, it is forced to make a higher offer. In this case, the lowest wage that would induce some workers to stay is to match the offer of the raider. Therefore, we can determine the equilibrium from the condition that the optimal $v_p(t)$ makes zero profits. The raider will lose money on workers that have ability $\eta < v_p(t)/\alpha(t)$ but will make positive profits from workers who have $\theta > 0$ and decide to leave despite the counter-offer of the firm. Thus, the zero-profit condition implies that

$$[1 - \lambda[1 - G(w_R(t) - v_p(t))]] \int_0^{w_R(t)/\alpha(t)} [\alpha(t)\eta - v_p(t)]dF(\eta)$$
$$+\lambda[1 - G(w_R(t) - v_p(t))][\alpha(t)\bar{\eta} - v_p(t)] = 0$$

[9]However, this strategy has costs if all of the potential raiders 'tremble'. Thus we have not shown that this equilibrium is Trembling Hand Perfect.

Now suppose $v_p(t) < w^*(t)$, then a raider could increase its offer and will not get any more lemons, since it is already receiving all $\eta < w_R(t)/\alpha(t)$, but will have a higher probability of getting workers who have high productivity and $\theta > 0$ (a formal proof would follow Lemma 1). Thus we can concentrate on outcomes that have $v_p(t) = w_R(t) \geq w^*(t)$. This is sufficient to prove the following result:

Proposition 5. *With poaching offers made at any time, there exists a unique equilibrium training level, $t^r > 0$. Denote the set of equilibrium training levels without poaching by T^{np} and the set of equilibrium outside wage functions by V^{np}, then the equilibrium training level with poaching, $t^r \leq \min T^{np}$ and the equilibrium outside wage offer $v^p(t) \geq \max V^{np}$.*

Proof: See the appendix.

To reiterate, the basic idea is that the firm is still able to retain some of the high ability workers, and given Assumption 1, this makes training worthwhile. However, the more aggressive raids tend to reduce the level of training.

This analysis also illustrates that our model has quite rich predictions and generates positive levels of training in different institutional settings. Moreover, it suggests a crucial role for the organization of the labor market. In a market where institutions make poaching more difficult (such as works councils in Germany which enforce consistent rules in hiring and promotions), we predict that training levels should be higher than in a labor market where raids are not restricted (e.g. the U.S.).

3.1.3. Some Further Issues

While we have addressed the issue of raids, there is another question that can be posed. Workers who are of high ability but unhappy at their current job could quit after the firm's retention decision and thus signal to the market that they were not actually laid off. This is not allowed in our model. It has to be noted that our model is too simplistic in that learning is instantaneous and complete. A more realistic but less transparent model would have abilities changing over time and new information arriving constantly regarding the productivity of the worker. In such a model with continuous learning, quits at any point in time cannot be distinguished from layoffs immediately after the arrival of bad news. A model with a continuous flow of information would also be more useful as the current model has the extreme assumption that all relevant information about the worker

is learned at the end of the first period and firms have no opportunity to fine tune training to the workers' needs and ability. Although with continuous learning the firm can vary the amount of training it gives, such a model would still imply that training will not be a one-to-one function of ability (otherwise the informational advantage of the firm would completely disappear).

3.2. Equilibrium with Verifiable Training

The assumption we have used so far has been that the firm offers only a training wage to attract workers and decides unilaterally about the training level. We now allow firms to offer a contract $[W, t]$ to attract workers to their training programs. The optimal contract is found by maximizing the expected utility of a worker subject to a zero-profit condition and the liquidity constraint of the worker. Thus the problem reduces to

$$\max_{W,t} (1-\lambda)[1-F(\hat{\eta}^*(t))]w^*(t) + \lambda[1-F(\hat{\eta}^*(t))][1-G(w^*(t)-v(t))]v(t)$$
$$+[\lambda + (1-\lambda)F(\hat{\eta}^*(t))] \int_0^{w^*(t)-v(t)} [w^*(t)-\theta]dG(\theta) \tag{3.1}$$

subject to

$$\begin{aligned} W &= [1-q[w^*(t^*), v(t^*)]] \int_{\hat{\eta}^*(t^*)}^{\infty} [\alpha(t^*)\eta - w^*(t^*)]dF(\eta) - t^* \\ W &\geq 0 \end{aligned} \tag{3.2}$$

where $w^*(t)$ and $\hat{\eta}^*(t)$ are determined through (2.4) and (2.5) by the maximization of the firm and $v(t)$ is determined by Bayes' rule by the outside firms as described by (2.8). The resulting allocation is summarized in the next proposition.

Proposition 6. *Generically, there exists a unique equilibrium level of training t^*. This training level is such that $t^c > t^* > \max T^{np}$, where T^{np} is the set of equilibrium training levels of section 2.*

Proof: See the appendix.

When the level of training is verifiable, the multiplicity of equilibria disappears even without raids. In addition, the amount of training is higher. The intuition

for both results is related. First recall that with training unilaterally chosen by the firm as in Section 2, each firm could believe that outside wages would be high, thus ex post monopsony power limited, and it would want to offer little training. Can we have the same hold in an equilibrium with verifiable training? The answer is no. Since the zero-profit condition holds in the second-hand labor market, workers will anticipate that in an equilibrium with high quits, they will still benefit from high training. They will therefore be willing to assume some of the cost of training by accepting a contract with a lower wage W. Thus, even when it is believed that ex post monopsony power will be limited, offering high training is profitable for firms because such a package will attract workers at low starting wages. This also explains why firms give more training now than in section 2. The next question to ask is why this process does not take us to the first-best. The answer is instructive; even though the package specifies the training level, it does not commit the firm to a firing and subsequent wage policy. This implies that there will be some degree of adverse selection in the second-hand market, and as a result some workers with low values of θ will prefer to stay in their training firm, thus first-best will not be achieved. Moreover, as productivity and training are complements, higher training levels imply that ex post quit decisions will be more distorted, hence the equilibrium training level lower than t^c.

3.3. Relation to the Literature

Our model is related to a number of papers in the literatures on training and on wage determination in imperfect labor markets. Highlighting the differences will further clarify our main points. The first strand of the literature is the competitive models of training. These models start with the seminal work of Becker (1964) and Mincer (1974). As we show in the second section, these models are unable to generate general training in the absence of a significant firm-specific component in the apprenticeship process unless workers pay for the training. Thus it is apparent that the fundamental logic of our mechanism is rather different.

Regarding the German apprenticeship system, Oulton and Steedman (1994) have conjectured that training wages simply reflect productivity, so that the general human capital model explains the German case. This view is not widely shared, since various studies by the German government and by the European community have repeatedly found that large enterprises have significant net outlays during apprenticeship training (see Harhoff and Kane, 1994, and Soskice, 1994 for numbers and references). Training practitioners we have talked to, in-

cluding union representatives, confirm the interpretation that training is costly to firms. The question which remains, however, is whether firm-specific components of training are important enough to explain the existence of substantial general training as well. We therefore contrast our model to the specific human capital case in the empirical analysis below, to investigate which explanation seems more consistent with the evidence.

The second group of papers includes the studies of adverse selection and raids in labor markets by Greenwald (1986), Waldman (1984), and Lazear (1986). While Greenwald noted the link between the superior information of the firm and the profitability of general training, our paper is the first to formalize this argument. In fact, we show that such superior information will not always lead to general training but will do so when training and unobserved ability are complements. Our model also goes beyond these previous treatments of the adverse selection problem, in allowing raids, in endogenizing the quit decisions of workers, and in deriving a possible multiplicity of equilibria.[10]

Another paper that is related both theoretically and empirically is Gibbons and Katz (1991). They too construct an asymmetric information model and test the implications of this model using the wages of displaced workers and quitters. The first difference is that Gibbons and Katz do not consider training, nor do they analyze endogenous quits and raids. In addition, in contrast to theirs, our test, described in more detail in the following subsection, involves a third group, so we can draw a sharper distinction between an adverse selection model and models with pure heterogeneity.

There are other mechanisms which lead to monopsony power of the firm without deriving this from asymmetric information. In this case too, the monopsony power of the firm would imply that it is willing to pay part of the general training cost. This result is embedded in the more general model of Acemoglu (1995), where the monopsony power arises from imperfect labor markets and wage bargaining. Acemoglu's paper focuses on the interactions of training and innovation decisions, while the basic intuition can be traced back to Grout (1984). Acemoglu (1995) also shares with this paper the emphasis in the quality composition of the labor force in determining training (and innovation) decisions.

[10]There are a number of other papers that point out the possibility of multiple equilibria in the presence of adverse selection. See, for instance, Chang and Wang (1994), Abe (1994), and Pendergrast (1992). None of these papers deal with the case of general training that is observed by all parties (a realistic assumption especially for the German system), nor do they analyze how the result of multiplicity is affected by other features of the model, including the contractual structure and the possibility of raids.

4. Deriving Testable Implications and Empirical Strategy

We now draw implications from the model which we will test below with data from Germany. We also contrast the implications of our theory to a specific human capital model. There are three main reasons why apprentices leave their firm in Germany: (i) their employer does not offer a permanent contract after the apprenticeship, (ii) workers quit voluntarily to find other employment; (iii) men are drafted to do military service. We assume that firms are unable to distinguish between voluntary quits and layoffs, but can see if a worker comes back from military service. The fact that there is another group of workers which separates for military service does not change any of our theoretical analysis. In our data we can distinguish workers who stayed with their training firm initially, those who left for military service and then work for a different employer, and workers who leave for other reasons. We do not know whether workers who stay with their apprenticeship firm served in the military. The model implies the following:

1. Stayers earn more than laid off workers and voluntary quitters. That is $w^*(t) > v(t)$. It is easy to see that this has to be true. If $w^*(t) < v(t)$, stayers would quit even if they did not receive a negative taste shock and be better off in the second-hand market.

2. Military quitters leaving for largely exogenous reasons earn more than other quitters. That is $v_m(t) > v(t)$. The reason they earn more than quitters is that they are of higher average ability and both groups are paid their average marginal product in the second-hand market. For this to be true it is not necessary that there is no selection at all among military quitters. For example, training firms may want to rehire very high ability military quitters after they complete their service. If there is some cost to doing so —for instance, in the case that firms need to keep certain positions open for these workers—, training firms will be more selective in retaining workers who have to do military service than other workers. Denote the ability cut off for military quitters used by the training firm in equilibrium as $\hat{\eta}_m(t)$, then the average quality of the pool of military quitters in the second-hand market is

$$\frac{q[w^*(t),v(t)]\alpha(t)\bar{\eta} + [1-q[w^*(t),v(t)]]\int_0^{\hat{\eta}_m(t)}\alpha(t)\eta dF(\eta)}{q[w^*(t),v(t)] + [1-q(w^*(t),v(t)]]F(\hat{\eta}_m(t))}$$
$$> \frac{q[w^*(t),v(t)]\alpha(t)\bar{\eta} + [1-q[w^*(t),v(t)]]\int_0^{\hat{\eta}^*(t)}\alpha(t)\eta dF(\eta)}{q[w^*(t),v(t)] + [1-q(w^*(t),v(t)]]F(\hat{\eta}^*(t))}$$

where the inequality follows from $\hat{\eta}_m(t) > \hat{\eta}^*(t)$.

3. Military quitters may earn more or less than stayers, $v_m(t) >< w^*(t)$. The reason for this is that military quitters are freed from the adverse selection problem (at least to some degree). If the monopsony power of the training firm over stayers is large enough compared to the quality of the pool of military quitters, the equilibrium wage for stayers may be below the wage of workers separated for an exogenous reason (see for instance the numerical example in section 2.3.3).

4. A specific human capital model which could also generate training paid for by firms will imply $v_m(t) < w^*(t)$ because military quitters loose their specific human capital. Since our data are cross sectional with retrospective information on a worker's apprenticeship firm, many stayers have also separated from their training firm by the time of the survey. We therefore control for whether a worker is still with his apprenticeship firm when we observe earnings. While this has no bearing on the ranking of earnings among the three groups for our model, the specific human capital model implies

$$\begin{aligned} & E(v(t) \mid \text{separated from training firm}) \\ = & E(v_m(t) \mid \text{separated from training firm}) \\ = & E(w(t) \mid \text{separated from training firm}) \end{aligned}$$

Our empirical strategy to test these implications is therefore to run a wage regression controlling for worker characteristics at the time the apprenticeship starts (e.g. information on schooling and on the type of apprenticeship firm). Dummy variables for stayers and military quitters let us discern the wage differentials for these two groups relative to other quitters.

5. The Data

We use data from two cross-sections of the German "Qualification and Career Survey"[11] conducted in 1979 and in 1985/86. Both surveys sample approximately 30,000 employed workers of German nationality in the age group 15 to 65. Besides

[11] Qualifikation und Berufsverlauf, conducted by the German Federal Institute for Vocational Training (Bundesinstitut für Berufsbildung) and the Institute for Labor Market Research (Institut für Arbeitsmarkt- und Berufsforschung).

standard questions on demographics and the current job, the surveys contain rich information on job attributes and job content, the qualifications needed for the job and how they were obtained, and retrospective questions on a worker's career path, education, and training history. In particular, the survey contains information on the apprenticeship firm of a worker, and whether and why a respondent separated from that firm.

We restrict the samples we use to workers who went through an apprenticeship and did not obtain any higher school based education. Additional education after the apprenticeship would introduce another group of quitters which is not easily integrated into our model. To make sure that we look at a population which is likely to have completed both their apprenticeship and their military service we limit the sample to those between age 23 and 59. We only consider men working full time because women have very different career patterns and they often train for very different occupations than men. About 99 percent of men with the other sample characteristics work full time in Germany. We also eliminate workers whose apprenticeship was in the public sector or who work currently in the public sector or in the construction industry. Retention rules after the apprenticeship and pay determination in the public sector are likely to be governed by rules different from the private sector. The construction industry is the only sector in Germany were training is not financed voluntarily by firms. Instead, a levy is imposed on those firms who do not train apprentices and redistributed to those who do. This indicates that turnover in the construction industry may be sufficiently high that the mechanism we discuss in this paper does not support training by firms. Furthermore, we restrict the sample to those who left secondary school in 1948 or later. Most military quitters in the earlier period served in World War II, which may introduce survivorship or other biases. In addition, the environment in which firms operated before the war might have been different, since the industrial apprenticeship system had developed only since the turn of the century.

The earnings variable on the surveys is the gross monthly wage. Respondents in the 1979 survey were asked to report their earnings in 13 brackets and in the 1985/86 survey in 22 brackets. We assign each individual earnings equal to the bracket midpoint.[12] Because of the large number of brackets this is unlikely to introduce much more measurement error than is done by respondents' rounding

[12]The top bracket in 1979 was DM 5,000 or more which we assigned a value of DM 7,500 (2.9 percent of all sample respondents are in this bracket), in 1985/86 it was DM 15,000 or more which we assigned a value of DM 17,500 (0.6 percent of sample respondents). The results are not sensitive to the value assigned to the top bracket.

continuous amounts. We then convert the variable to an hourly wage by dividing by the number of weekly hours. We construct the standard variable for potential experience as age - schooling - 5. The number of years of schooling are derived from the survey information on the types of schools attended and degrees obtained following Krueger and Pischke (1995).

Our key variables of interest are related to the separation from the training firm after the apprenticeship. The survey asks when respondents left their training firm, and the reason for the separation. We only focus on separations immediately after the apprenticeship. Apprenticeship contracts in Germany are generally fixed term, and a new contract has to be written for further employment. Typically, a training firm will make employment offers to most of its apprentices but it is under no legal obligation to do so. As reasons for a separation, the 1979 survey allowed six responses: left for military service, laid off or quit, could obtain a higher income elsewhere, better working conditions elsewhere, obtained additional training or education, and other. Multiple responses were allowed. In 1985 only three mutually exclusive answers were possible: military conscription, laid off by the firm, or quit voluntarily. We create two dummy variables from these questions. The variable *stayers* is used for those who stayed with their apprenticeship firm after the end of the apprenticeship, although these workers may have left the firm by the time of the survey. The variable *military quitters* is one for those who left their apprenticeship firm immediately and mention military service as the reason (even if other categories are also mentioned in the 1979 survey). The variable also includes conscientious objectors leaving for a non-military service.

We will think of the military quitters as separated from their firms due to exogenous reasons and we also assume that firms taking part in the secondary labor market for trained workers will realize this. Since this is our key identifying assumption, the possibility of systematic selection into the military deserves some discussion. Military service has been mandatory for all able bodied German males since 1957, but the fraction of a cohort actually drafted has varied over time. The military conducts a physical exam, but no other tests to determine draft eligibility. There are also other temporary exemptions, mainly to allow young men to complete an education in process. This will typically not apply to youths leaving school at age 15 or 16 beyond the time when they complete an apprenticeship of three years. Since the 1960s, about 60 to 75 percent of a cohort have served (Ullrich, 1984). The pure mechanics of the draft system in Germany severely limit the military's ability to select draftees.

In order to assess whether our results may be biased due to selection into

or experience from military service we also use data from the first wave of the German Socioeconomic Panel (SOEP). This dataset has retrospective questions on the employment history of all workers from age 16. In particular, we know whether any respondent served in the military at any particular year of their life. Using these retrospective data we construct a dummy variable for everybody who ever served. We select a sample analogous to the Qualification and Career Surveys. However, unlike the samples from the Qualification and Career Surveys, the SOEP sample includes respondents who might have returned to higher education after the apprenticeship and those who did their apprenticeship in the public sector but work in the private sector at the time of the 1984 survey. These exclusions are not possible in the SOEP. The sample includes 526 observations.

6. Empirical Results

6.1. Summary Statistics

Table 1 gives summary statistics for the two data sets, for the full samples as well as for stayers and military quitters separately. Most apprentices leave secondary school after 9th or 10th grade, only few obtain additional schooling before the apprenticeship although the incidence of higher schooling increases over time. Military quitters are more likely to have a tenth grade education. This may result from the fact that many teenagers leaving school after grade 9 may not have reached draft age (18) when they complete their apprenticeship. We therefore control for this difference in schooling in the regressions below. Slightly less than 10 percent of sample members go on to obtain a qualification as master craftsmen. This entails refining the skills in their trade, and allows them to set up their own crafts shop and to train apprentices themselves. Military quitters are younger because significant numbers of young men were drafted starting only in the 1960s. We constructed two measures of work related disabilities. Severe disabilities refer to an officially recognized disability rated at 50 percent or more of a regular worker's capacity. Any other reported disability is classified as less severe. Military quitters tend to be more able bodied, since the military screens on health, but the differences are slight.

About 80 percent of apprentices stay in their apprenticeship firm initially although only 25 to 30 percent remain there by the time of the survey.[13] The

[13]The 1979 survey failed to allow for an exhaustive set of answers to this question so that the classification of who is still working in their apprenticeship firm may be error ridden. It

fraction of workers still in their apprenticeship firm is slightly higher among stayers, but not dramatically so. Military quitters make up only 3 to 5 percent of the sample, much fewer than the fraction of young men in Germany who served when the military draft was in place after 1957. This presumably reflects the fact that many workers stay with their firm after their apprenticeship and only then interrupt their career for military service or return to their previous employer even after military service. This does not invalidate our test, as discussed in section 4. Nevertheless, as figures 1 and 2 show, the fraction of military quitters varies over time (i.e. by graduation year from the apprenticeship) in the fashion expected from aggregate draft rates. The fraction of military quitters increases from the mid-1950s and stabilizes around the mid-1960s. Note, however, that the estimates in these figures are rather noisy due to the small cell sizes.

A large fraction of apprenticeships takes place in the crafts sector and in smaller firms, although manufacturing is also a large provider of training. Many apprentices move out of the crafts sector; while it trains more than 40 percent of our samples, it only employs a little more than 20 percent. Most of the movement is into services and other industries. This category include professions such as lawyers, accountants, and physicians, who train paralegals, and accounting and physician's aides. The distribution across job types is rather similar across groups. About 20 percent in 1979 and 15 percent in 1985/86 of apprenticeship trained employees work in relatively unskilled blue or white collar positions. 7 percent of the 1979 sample reports reaching an executive position compared with only 1 percent in 1985/86. The relevant question in the later survey was more precise in stating that this group only refers to employees with extensive managerial discretion. Military quitters report that they use slightly less of their apprenticeship skill although even this difference practically vanishes in the later survey. More than 65 percent of workers use their apprenticeship skills "much" or "a lot."

6.2. Basic Results

Stayers and military quitters differ in many observable dimensions, including some that were clearly discernible at the beginning of the apprenticeship. Since there is likely to be a good deal of selection into better and worse apprenticeships among school leavers (Soskice, 1994, argues that this is the case), it is important to control

allows responses for those who left within the first five years after their apprenticeship and those who stayed but no category for those who left after five years is allowed. Item nonresponse is no higher than in the 1985/86 survey so that the members of the missing category must be distributed in the remaining cells.

for these observable attributes in our wage regressions. On the other hand, we do not want to control for any attributes of the jobs workers obtained after their apprenticeship since these will be related to worker ability and selection, which is part of the outcome we are interested in. We therefore present wage regressions in table 2, which include the number of years of primary and secondary school, a quartic in potential experience, the sector of the apprenticeship firm, two dummies for disability status, and a dummy for workers who are still in their apprenticeship firm. The last variable is the only one which potentially contains information on post-apprenticeship outcomes. It turns out to be relatively unimportant and we also present results without this variable.

Our main result is that both stayers and military quitters earn higher wages than those who left their apprenticeship firm for other reasons. Only the coefficient for stayers is significant in the 1985/86 regressions, while the others are not. Nevertheless, the consistency of the pattern across surveys makes it more likely that the results are not merely the product of sampling variability. In fact, restricting the coefficients on the two dummies to be the same across surveys yields a value of 0.038 with a standard error of 0.018 for military quitters and 0.032 for stayers (standard error 0.012). Both are significantly different from zero below the 2 percent level. The fact that military quitters may actually earn more than stayers is in support of one of our key predictions. Recall that military quitters have higher ability than other quitters but lower ability than stayers. However, once they separate military quitters can obtain their marginal product while the incumbent firm extracts rents from the stayers. If these rents are large enough, they may outweigh the quality advantage of the stayer group (see the numerical example above). Thus, the test of whether military quitters earn higher or as high wages as stayers is a challenging test of our predictions.[14]

An alternative interpretation of our results would be that the military positively selects draftees. We argued already above that we find it unlikely that such selection is important since the military only screens on health. Our controls for work related disabilities should pick up differences in health status of veterans and other workers. Another possibility is that military service imparts skills or attitudes which are valued in the civilian labor market. Again, we find this un-

[14]Notice, however, that stayers will mostly have obtained other jobs by the time of the survey (and we control separately for still being in the apprenticeship firm), so that these rents do not accrue to the apprenticeship firm anymore. The adverse selection mechanism which drives our model of training is likely to continue to be important later on in workers' careers. This is true in particular if ability evolves over time and more information accrues to the incumbent firm.

likely to be the case for skilled workers in Germany. Imbens and van der Klaauw (1995) find that mandatory military service in the Netherlands lowers earnings by about 5 percent for those who served, which implies veterans lose approximately the returns to experience during the time of service. This suggests that there is little evidence of military service raising earnings in civilian jobs, for example through training.

To provide more direct evidence on this question we ran a wage regression for a similar sample of men from the 1984 wave of the Socioeconomic Panel. This dataset allows us to include a dummy variable for everybody who ever served in the military[15]. If draftees are positively selected or if military experience is valuable in the civilian labor market then military service should be associated with a positive wage differential. Instead we find that service actually leads to slightly lower earnings, although this effect is not significant. This indicates that our findings for military quitters are very unlikely to be due to selection or training in the military, thus leaving our explanation as the likeliest candidate.

6.3. Robustness Checks

We investigate the robustness of our results next and extend our findings in some other directions. Table 3 reports the coefficients on the stayer and military quitter dummies for a number of different samples and specifications, making modifications to the baseline one at a time. The 1979 survey allows us to control for secondary school grades in math and German. This lowers the relevant coefficients slightly. Dropping the dummy variable indicating whether a worker is still with his apprenticeship firm leads to slightly lower coefficients on the stayer variable.

The same is true when we eliminate self-employed from the sample, a group whose income is typically less accurately reported. We also investigate the effect of excluding those with 12 or 13 years of education. This group often goes on to additional schooling after the apprenticeship. We have eliminated anybody who obtains further academic credentials from our sample already, because this introduces an additional group of quitters after the apprenticeship which does not easily fit in our simple model. This may leave a curious selection of 12th and 13th graders who remain in the sample. Again the coefficients are slightly lower when this group is dropped.

[15]This dataset does not allow us to construct some of the other controls about the type of apprenticeship because such data are not collected on the survey.

To ensure that our military quitter dummy does not pick up any nonlinearities in age or experience not already controlled for, we limit the sample to those who were born after 1938. These are the cohorts at risk of being drafted for military service. In addition, instead of controlling for a quartic in experience, we put in separate dummies for each age of the respondents in the sample. The coefficient on the military quitter dummy is slightly higher. The next row takes the opposite step and drops the sample restriction that respondents finished school after 1948. This adds more observations particularly to the 1979 survey, many of whom completed their apprenticeship during the 1930s and 40s. This increases the diversity of military quitters by including World War II veterans. It turns out that this inclusion lowers the coefficients on the military quitter dummy by about 4 log points in the 1979 survey. There is no change in the later data set. The most dramatic and inconsistent change occurs when we limit the sample to apprenticeships in the manufacturing industry. This is the sector where training is of the highest quality and the net costs borne by the firms are the largest. In fact, in table 2 we found an 8 to 10 percent wage differential for training in manufacturing over the crafts sector. Thus, we might expect to find the most pronounced support for our model when we limit the sample to manufacturing. Instead, the stayer coefficient goes to zero in the 1979 sample, and the military quitter coefficient does the same in the later data. However, we are eliminating two thirds of the sample with this restriction and standard errors more than double so that these results may simply not be very reliable.

In summary, the results in tables 2 and 3 confirm the predictions of our model. Workers who stay with their firm after the apprenticeship earn higher wages than those who quit or are laid off. Military quitters also earn higher wages than other quitters because they are separated from their firm for an exogenous reason and should therefore be of higher quality than other quitters. In fact, there earnings are as high or higher as those of stayers.

6.4. Can There be Alternative Interpretations?

It is important to point out that these results are not easily reconciled with alternative explanations. One objection might be that there are ability differences between stayers, quitters, and military quitters because of selection, but these differences are observable to everyone. Notice, however, that this explanation suggests that military quitters should earn strictly less than stayers, who are the most able group.

A simpler model of training might be based on specific human capital. If a large enough part of apprenticeship skills is firm-specific, the training firm may be able to recoup its investment even in the general human capital component of the apprenticeship. We include a dummy for everybody still with their apprenticeship firm in the wage regressions in table 2. In a specific human capital model, the coefficient on this variable should be positive. On the other hand, the model says nothing about the earnings of leavers depending on the timing of the separation. Thus, we would expect zero coefficients on the stayer and military quitter dummies. The empirical results show exactly the opposite pattern.

An additional piece of evidence for our interpretation of the data is provided by Pischke's (1994) analysis of further training in Germany. He finds that the returns to such training are basically zero. This result does not differ much between workers who stay with the firm which provided the training and workers who leave the training firm. In terms of the point estimates, the returns for stayers are even slightly negative. This indicates that higher wages for stayers after firm sponsored training are unique to apprenticeship training. This is, of course, the period when the adverse selection problem is most severe because there is very little information about worker abilities which is observable to outsiders.

6.5. Additional Evidence

Our model gives a clear ranking of the relative abilities of the three groups. Stayers should be the most able, military quitters ought to fall in the middle, and other quitters should have the lowest ability. Therefore, a variable related to ability would let us assess this prediction. Such a measure should not be publicly observable before apprenticeship training takes place, because the quality and extent of training itself may depend on the observed ability of secondary school leavers as they get sorted into different firms.

Entorf, Gollac, and Kramarz (1995) argue that computers are used by more able workers. Table 4 reports results from linear probability models for using a computer on the job on the same controls as in table 2. Stayers are significantly more likely to use computers in the 1985/86 survey while the incidence of computer use is the same among military quitters and other quitters.

Another variable possibly correlated with ability is promotion. The clearest indication of advancement is the position of a shop floor supervisor among blue collar worker. This position is usually associated with further training and often implies that the worker may be training apprentices himself. Stayers are more

likely to reach this position while the effect is negative for military quitters. We present results both for the full sample used previously and for blue-collar workers only since shop floor supervisors are classified as blue collar. The results are slightly more favorable in the restricted sample.

Traditionally, the position of shop floor supervisor was associated with master craftsman training but this may be less the case in large manufacturing firms where little is gained from this formal qualification. Nevertheless, master craftsmen training may also be an indicator of ability. Among blue collar workers both stayers and military quitters are more likely to undertake such training than other quitters in 1979 and the opposite is true in 1985/86. The effects are always stronger for military quitters.

Overall, the evidence on computer use, promotion to shop floor supervisor, and master craftsmen training seems to indicate that stayers are of higher ability than military quitters which supports our hypothesis that military quitters earn wages which are higher than (or as high as) stayers because they are spared of the monopsony power of their apprenticeship firms. On the other hand, military quitters do not seem to differ much from other quitters which is not in line with our predictions. However, it should be borne in mind that the group of military quitters is small, so the results for this group are estimated very imprecisely.

7. Conclusion

This paper has presented a new theory of training. We believe this model has some attractive features relative to the standard Beckerian theory as it predicts that workers should not pay for their general training. We also argued that this theory fits the institutional structure of the German labor market, and thus it has particular relevance to Germany. Our model also sheds light on the very different patterns of training observed in Germany and the U.S. Despite the highly stylized nature of the model, we can make the following simple observations.

1. In terms of our model, we can think of the U.S. at a low training and high quit equilibrium and Germany at a low quit, high training equilibrium. The informal claim made by Blinder and Krueger (1991) that turnover should be negatively related to the level of training is formalized in our proposition 2. Importantly, both turnover and training are endogenous in our model.

2. The institutional differences, especially the role of works councils in Germany, make raids on workers in employment with other firms much harder

(see for example Abraham and Houseman, 1993). This suggests that Germany should have more training than the U.S. as formalized in propositions 4 and 5.

3. Information regarding the ability of the worker may be more important in Germany than the U.S. because of high firing costs at the later stages of a worker's career (i.e. after the apprenticeship). Since a firm can easily lay off its workers in the U.S., the cost of hiring a low quality worker is not enormous and thus adverse selection can be less of a problem. This observation also helps in thinking about the institutional differences between the labor markets of the U.S. and Germany in endogenous terms. High firing costs in Germany combined with the training system and the low quit equilibrium may not be too distortionary, while the same level of firing costs in the U.S. without any other changes in the labor market may be more damaging. This may help explain the good performance of the German economy during the post-war period despite the fact that the high degree of labor market regulation might have hampered the optimal allocation of resources.

4. Given the lower ability of firms to exploit internal training opportunities, our model also suggests that school based training financed by workers themselves should be more important in the U.S. This is true even if vocational schools are not as cost-effective as the apprenticeship programs of German firms.

5. When adverse selection in the labor market is important, this can explain firm sponsored training at all stages of a worker's career. Pischke (1994) reports evidence that further training in Germany does not lead to higher wages. This is true even though this training seems to be general: many workers reported receiving a written certificate which they would present when applying for a new job. Workers who received further training also report that it enables them to do their jobs better, so the training is productive. If similar levels of further training are given to workers of different abilities, adverse selection will prevent that the workers can capitalize on their additional skills in the market. This may explain why the incidence of further in-house training is higher in Germany than in the U.S.

6. Any explanation for training financed by firms will have to rely on some imperfection in the labor market. If labor markets are imperfect, compar-

isons of the U.S. and German training systems which rely on wages to make inferences about productivity might be misleading. For instance, Harhoff and Kane (1994) and borrowing from them, Heckman (1994), point out that there may be no market failure preventing training in the U.S. since wages grow at the same rate over the life-cycle in the two countries. In the model of this paper compensation is front loaded. Front loading is more extreme in equilibria with more monopsony power and thus with more training. Therefore, similar wage profiles in the U.S. and Germany are easily consistent with more human capital accumulation in Germany.

7. Our model is also consistent with the historical evolution of the apprenticeship system in Germany. From the middle ages till the 19th century apprentices paid a fee ("Lehrgeld") to their training firm. At this time, apprentices would typically choose self-employment after training in becoming journeymen or setting up their own crafts shop. Thus, adverse selection problems in hiring were unimportant. This changed when large enterprises started training and hiring skilled workers at the beginning of this century. By the time of World War II, positive training wages had replaced the fees paid by apprentices.

References

[1] Abe, Yukiko (1994) "Specific capital, adverse selection and turnover: A comparison of the United States and Japan," *Journal of the Japanese and International Economies* 8, 272-294

[2] Abraham, Katherine G. and Susan N. Houseman (1993) *Job Security in America: Lessons from Germany.* Washington, DC: Brooking Institution

[3] Acemoglu, Daron (1995) "Training and innovation in an imperfect labor market," mimeographed, MIT

[4] Becker, Gary (1964) *Human capital.* Chicago: The University of Chicago Press

[5] Bers, Lipman (1957) *Topology.* New York: New York Univeristy

[6] Blinder, Alan S. and Alan B. Krueger (1991) "International differences in turnover: A comparative study with emphasis on the U.S. and Japan," mimeographed, Princeton University

[7] Burdett, Kenneth (1979) "A useful restriction on the offer distribution in job search models," in Gunnar Eliasson, Bertil Holmlund, and Frank P. Stafford (eds.) *Studies in labor market behavior: Sweden and the United States.* Stockholm: IUI Conference Report

[8] Chang, Chun and Yijiang Wang (1994) "A framework for understanding differences in labor turnover and human capital investments," mimeographed, University of Minnesota

[9] Cho, In-Koo and David M. Kreps (1987) "Signalling games and stable equilibria," *Quarterly Journal of Economics* 102, 179-222

[10] Entorf, Horst, Michel Gollac, and Francis Kramarz (1995) "New technologies, wages, and worker selection," mimeographed, University of Mannheim

[11] Gibbons, Robert and Lawrence F. Katz (1991) "Layoffs and lemons," *Journal of Labor Economics* 9, 351-380

[12] Greenwald, Bruce C. (1986) "Adverse selection in the labor market," *Review of Economic Studies* 53, 325-347

[13] Grout, Paul (1984) "Investment and wages in the absence of binding contracts: A Nash Bargaining approach," *Econometrica* 52, 449-460

[14] Harhoff, Dietmar and Thomas J. Kane (1994) "Financing apprenticeship training: Evidence from Germany" NBER Working Paper No. 4557

[15] Heckman, James (1993) "Assessing Clinton's program on job training, workfare, and education in the workplace" NBER Working Paper No. 4428

[16] Imbens, Guido, and Wilbert van der Klaauw (1995) "Evaluating the cost of conscription in the Netherlands," *Journal of Business and Economic Statistics* 13, 207-215

[17] Krueger, Alan B. and Jörn-Steffen Pischke (1995) "A comparative analysis of east and west German labor markets: before and after unification," in: Richard B. Freeman and Lawrence F. Katz (eds.) *Differences and changes in wage structures.* NBER Series in Comparative Labor Markets, Chicago: University of Chicago Press, 405-445

[18] Lazear, Edward P. (1986) "Raids and offer matching," in: Ronald Ehrenberg (ed.) *Research in labor economics* 8, 141-165

[19] Lynch, Lisa M. (1994) "Introduction" in Lisa M. Lynch (ed.) *Training and the private sector. International comparisons.* NBER Series in Comparative Labor Markets, Chicago: University of Chicago Press, 1-24

[20] Mincer, Jacob (1974) *Schooling, experience, and earnings*, New York: Columbia University Press

[21] Noll, I., Ursula Beicht, Georg Boll, Wilfried Malcher, and Susanne Wiederhold-Fritz (1983) *Nettokosten der betrieblichen Berufsausbildung.* Schriften zur Berufsbildungsforschung 63, Berlin: Bundesinstitut für Berufsbildungsforschung

[22] Osterman, Paul (1993) "How Common is Workplace Transformation and How Can We Explain Who Adopts It?" mimeographed, Sloan School, MIT.

[23] Oulton, Nicholas and Hilary Steedman (1994) "The British system of youth training: a comparison with Germany" in Lisa M. Lynch (ed.) *Training and the private sector. International comparisons.* NBER Series in Comparative Labor Markets, Chicago: University of Chicago Press, 61-76

[24] Pendergast, Canice (1992) "Career Development and Specific Human Capital Collection," *Journal of the Japanese and International Economies* 6, 207-227

[25] Pischke, Jörn-Steffen (1994) "Continuous training in Germany," mimeographed, MIT

[26] Soskice, David (1994) "Reconciling markets and institutions: the German apprenticeship system" in Lisa M. Lynch (ed.) *Training and the private sector. International comparisons.* NBER Series in Comparative Labor Markets, Chicago: University of Chicago Press, 25-60

[27] Stokey, Nancy L. and Robert E. Lucas (1989) *Recursive Methods in Economic Dynamics.* Cambridge, MA: Harvard University Press

[28] Topel, Robert H. and Michael P. Ward (1992);"Job Mobility and the Careers of Young Men" *Quarterly Journal of Economics* Vol 107, pp 439-480.

[29] Ullrich, A. F. (1984) *Die Bundeswehr. Eine Diskussionsgrundlage.* Dülmen, Germany: Laumann Verlag

[30] Waldman, Michael (1984) "Job assignments, signalling, and efficiency," *Rand Journal of Economics* 15, 255-267

8. Appendix: Proofs of Propositions and Lemmata

Proposition 1: (i) First ignore t^* and W^* and look at the determination of the three equilibrium functions $w(t)$, $v(t)$ and $\hat{\eta}(t)$ through the system defined by (2.4), (2.5) and (2.8). Note that $v(t)$ is determined in equilibrium from (2.8) and $w(t)$ and $\hat{\eta}(t)$ are determined by maximizing (2.3). Also note that as $t \to \bar{t}, \alpha'(t) \to 0$, throughout, we can restrict attention to a closed and bounded space whereby $w(t)$ and $v(t)$ lie in $[0, \alpha(t)\eta^{\max}]$ and $\hat{\eta}(t)$ lies in $[0, \eta^{\max}]$. Finally, to start with suppose that $v(t)$ is differentiable. Then, (2.3) is straightforwardly seen to be differentiable. Thus if there is a unique optimizer $w(t)$ and $\hat{\eta}(t)$ for every t, then $w(t)$ and $\hat{\eta}(t)$ as functions would be continuous and differentiable.

To check the uniqueness of the $w(t)$ and $\hat{\eta}(t)$, we need to study the concavity of (2.3). We first substitute for $\hat{\eta}(t) = w(t)/\alpha(t)$ and obtain the profit function of an apprenticeship firm as:

$$\max_{w(t)} \Pi = [1 - q[w(t), v(t)]] \int_{\frac{w(t)}{\alpha(t)}}^{\infty} [\alpha(t)\eta - w(t)] dF(\eta)$$

To ensure the uniqueness of the function $w(t)$ it is sufficient to have $\log \Pi$ concave.[16] Therefore we need to check that

$$\log \Pi = \log[1 - q[w(t), v(t)]] + \log \int_{\frac{w(t)}{\alpha(t)}}^{\infty} [\alpha(t)\eta - w(t)] dF(\eta)$$

is concave. Now note that $q[w(t), v(t)] = 1 - \lambda + \lambda G[w(t) - v(t)]$ and therefore log-concavity of G ensures that $\log(1 - q)$ is concave. Next, Assumption 3, that $\int_x \eta dF(\eta)$ is log-concave ensures that the second term is concave. Thus the second order conditions for a maximum are satisfied whenever the first-order conditions hold. Hence $w(t)$ and $\hat{\eta}(t)$ are unique. Thus we have $w(t) = \Omega(v(t))$ where Ω is a continuous and differentiable operator. Therefore, to establish the existence of equilibrium is to show that the following functional equation has a fixed point:

[16] Note that if $\log H(x)$ is concave then $\frac{H''(x)}{H(x)} - \frac{H'(x)^2}{H(x)^2}$ and if we are evaluating this at a stationary point, $H'(x) = 0$, then the second term disappears.

$$v(t) = P[v(t),t] \equiv \frac{q[\Omega(v(t)),v(t)]\alpha(t)\bar{\eta} + [1-q[\Omega(v(t)),v(t)]]\int_0^{\frac{\Omega(v(t))}{\alpha(t)}} \alpha(t)\eta dF(\eta)}{q[\Omega(v(t)),v(t)] + [1-q(\Omega(v(t)),v(t)]]F\left(\frac{\Omega(v(t))}{\alpha(t)}\right)} \tag{8.1}$$

To establish the existence of a fixed point to (8.1), we will make use of Schauder's Fixed Point Theorem (see Stokey and Lucas, 1989 or Bers, 1957). To do this we need to show that

1. P maps from a closed and bounded subspace of $V \subset C[X]$ into V itself where X is a metric space over which the functions v are defined.

2. P is a continuous operator over V.

3. The family $P[v(t),t]$ is equicontinuous for every $v \in V$. (Or alternatively that V is a compact and convex Banach Space with the chosen topology).

We will establish these features using the sup-norm as our metric.

1. First take $V \subset C[0, \bar{\eta}\alpha(\bar{t})]$. Clearly, no wage greater than $\bar{\eta}\alpha(\bar{t})$ will paid to a second-hand worker and also no wage less than 0. Second, limit attention to the set of functions that are differentiable with bounded derivative (i.e. those that satisfy a Lipschitz condition). This subspace is closed and bounded. Moreover, since P is continuous and differentiable with a bounded derivative (continuity and differentiability will be shown next, that P has bounded derivative, it suffices to look at $\partial P/\partial v$), it maps V into V itself, that is $P[v(t),t]$ will also satisfy a Lipschitz condition.

2. It is sufficient to show that P is differentiable. This follows as $v(.)$, $\Omega(.)$, $\alpha(t)$ are all smooth and $[0, \bar{\eta}\alpha(\bar{t})]$ is closed and bounded.

3. Take t and s such that $|s-t| < \delta$, we will show that there exists $\epsilon > 0$ such that $|P[v(s),s] - P[v(t),t| < \epsilon$ for all $v \in V$. Thus, we need to show that:

$$\left| \frac{q[\Omega(v(s)),v(s)]\alpha(s)\bar{\eta}+[1-q[\Omega(v(s)),v(s)]]\int_0^{\frac{\Omega(v(s))}{\alpha(s)}} \alpha(s)\eta dF(\eta)}{q[\Omega(v(s)),v(s)]+[1-q(\Omega(v(s)),v(s)]]F\left(\frac{\Omega(v(s))}{\alpha(s)}\right)} - \frac{q[\Omega(v(t)),v(t)]\alpha(t)\bar{\eta}+[1-q[\Omega(v(t)),v(t)]]\int_0^{\frac{\Omega(v(t))}{\alpha(t)}} \alpha(t)\eta dF(\eta)}{q[\Omega(v(t)),v(t)]+[1-q(\Omega(v(t)),v(t)]]F\left(\frac{\Omega(v(t))}{\alpha(t)}\right)} \right| < \epsilon$$

Now denote the numerator by $N_v(.)$ and the denominator by $D_v(.)$ as conditioned on the function v, then the problem can be written as

$$\left|\frac{N_v(s)}{D_v(s)} - \frac{N_v(t)}{D_v(t)}\right| < \epsilon$$

Since v, Ω and α are all continuous and defined over a compact set [the set of $[0, \bar{\eta}\alpha(\bar{t})]$], they are uniformly continuous, thus conditional on v, $\exists \epsilon_v > 0$ such that

$$\left|\frac{N_v(s)}{D_v(s)} - \frac{N_v(t)}{D_v(t)}\right| < \epsilon_v$$

Now since all $v \in V$ have bounded derivatives, there exists $\bar{\epsilon} > \epsilon_v$ for all $v \in V$, thus let $\epsilon = \bar{\epsilon}$. Thus the family $P[v(t), t]$ is equicontinuous and a fixed-point exists. Since V is the space of differentiable functions, the equilibrium functions $w(t), \hat{\eta}(t)$ and $v(t)$ are differentiable.

Given differentiability, it then follows that t^* that maximizes (2.6) and W^* that satisfies (2.7) can be found. QED

(ii) Now we establish some properties of the equilibrium. Suppose $t^* = 0$ is a solution. This implies $q[w^*(t^*), v(t^*)] = 1$ from Assumption 1 and (2.6), but this contradicts (2.5). Thus $t^* > 0$. Next suppose $t^* = t^c$. Then $q[w^*(t^*), v(t^*)] = 0$ and $\hat{\eta}^* = 0$ from (2.2) and (2.6), but then from (2.4), $w^* = 0$ and all $\theta > 0$ would leave, thus $q > 0$, a contradiction. Hence $t^* < t^c$. QED

Proposition 2: **(i)** The argument in the text establishes that the equilibrium quits and wages are given by the intersection of two upward-sloping schedules for a given t. Both are continuous by Assumption 2 and Proposition 1. Since $g(.)$ is unrestricted, we can always obtain multiple equilibria. Once we have multiple equilibria for a given training level t, then from (2.6) we can get the equilibrium levels of t.

(ii) First we show that $v^a > v^b$ is equivalent to $q^a > q^b$. Consider the first-order conditions of the maximization problem. Differentiating the definition of q we can see that

$$\frac{dq}{dv} = \frac{\partial q}{\partial v} + \frac{\partial q}{\partial w^*}\frac{dw^*}{dv}$$

where ∂v denotes a small change in v at given t. Next note that $\partial q/\partial v = g(w(t) - v(t))$ and $\partial q/\partial w^* = -g(w(t) - v(t))$. $\partial q/\partial w^*$ can be obtained by totally differentiating (2.5). This gives

$$\frac{dw^*}{dv} = \frac{\lambda g(.)(1-F(\hat{\eta})) - \lambda g'(.)\int_{\hat{\eta}}[\alpha\eta - w]dF(\eta)}{2\lambda g(.)(1-F(\hat{\eta})) - \lambda g'(.)\int_{\hat{\eta}}[\alpha\eta - w]dF(\eta) + \frac{1}{\alpha}\left[1-\lambda\left(1-G(.)\right)+\lambda g(.)\right]f(\hat{\eta})}$$

where $g(.)$ stands for $g(w-v)$ and similarly for $G(.)$ and $g'(.)$. Assumptions 2 and 3 ensure that both the numerator and the denominator are positive but the denominator is larger; thus $0 < dw^*/dv < 1$, which implies that $dq/dv > 0$ everywhere. Equivalence of $q^a > q^b$ and $t^a < t^b$ follows from (2.5). QED

Proposition 3: Take a sequence of distribution functions G_n that converge to G_∞ in the sup norm where G_∞ is dirac at 0. Let W_n be the sequence of training wages corresponding to the equilibrium with distribution function G_n. Then it is immediate that $W_\infty > 0$, since when G_∞ is degenerate $w^* = \hat{\eta} = v = q = 0$. Then $\exists\ n^* : \forall n > n^*, W_n > 0$. QED.

Lemma 1: All $v(t) \in V^{np}$ must be definition satisfy equation (2.8). Consider $v^s(t) = \max V^{np}$. Since this is the extremal equilibrium it is clear that if we increase $v(t)$ for a given t, the RHS of (2.8) will increase more than the LHS. Now, an outside wage function $v^p(t) > v^s(t)$, then all workers who would have quit or been laid-off would accept this offer. Therefore the per worker revenue of the firm is given by the RHS of (2.8) and thus this firm makes negative profits. Therefore, $v^p(t) \leq \max V^{np}$.QED

Proposition 4: Lemma 1 establishes that $v_p(t) \in V^{np}$. Now suppose that $v_p(t) < \max V^{np}$. This cannot be an equilibrium, because an outside firm can offer the wage function $\hat{v}(t) = \max V^{np} - \epsilon$ to quitters, for ϵ small enough more workers who are retained but unhappy would quit and the outside firm would make a positive profit. Thus a contradiction. QED

Proposition 5: From the argument in the text we can see that in equilibrium all workers with $\theta > 0$ will leave. Then the payoff of the training firm is

$$[1-\lambda]\int_{\frac{w_R}{\alpha}}^{\infty}[\alpha(t)\eta - w_R(t)]dF(\eta)$$

This expression cannot be at its maximum for $t = 0$ given Assumption 1, thus $t^r > 0$. Next we show that

$$[1-\lambda]\int_0^{\infty}\alpha'(t)\eta dF(\eta) < [1-q]\int_{\hat{\eta}}^{\infty}\alpha'(t)\eta dF(\eta)$$

By reducing w^* on the RHS of this equation we obtain the LHS (as $w^* \to 0$, we get $q \to \lambda$ and $\hat{\eta} \to 0$). Thus to demonstrate that the RHS is larger than the LHS, we need to show that the derivative of (2.6) with respect to w^* is positive. However, (2.6) is the derivative of (2.3), thus

$$\frac{\partial(2.6)}{\partial w^*} = \frac{\partial(2.5)}{\partial t} = -\frac{\partial q}{\partial w^*} \int_{\hat{\eta}}^{\infty} \alpha'(t)\eta dF(\eta) > 0.$$

Now obviously

$$[1-\lambda] \int_{\frac{w_R}{\alpha}}^{\infty} [\alpha'(t)\eta - w_R(t)] dF(\eta) < [1-\lambda] \int_0^{\infty} \alpha'(t)\eta dF(\eta)$$

and thus the result that $t^* < \min T^{np}$ follows. QED

Proposition 6: When firms offer a package $[t, W]$ to attract trainees, the problem reduces to maximizing the utility of the workers subject to a zero profit condition, (3.1) and (3.2). Next, given that the contract does not specify the layoff and second-period wages functions, the firm will always choose $\hat{\eta} > 0$, thus $v(t)$ will always be less than what the workers receive in the training firm, $w^*(t)$, thus the greater is t, the more distorted is the quit decision. Thus $t^* < t^c$. Since the equilibrium is now given as a result of a maximization problem rather than a fixed point, the equilibrium must assign a unique maximum value W to workers. Given the twice differentiability of the objective function, generic uniqueness of t^* follows. QED

Figure 1

Fraction of Military Quitters by Year, 1979 Survey

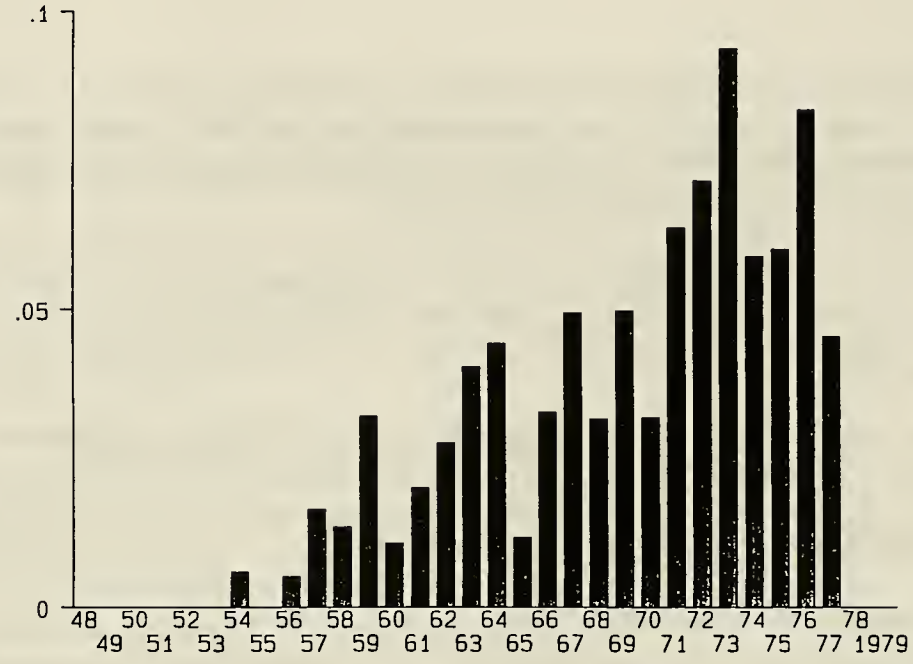


Figure 2

Fraction of Military Quitters by Year, 1985/86 Survey

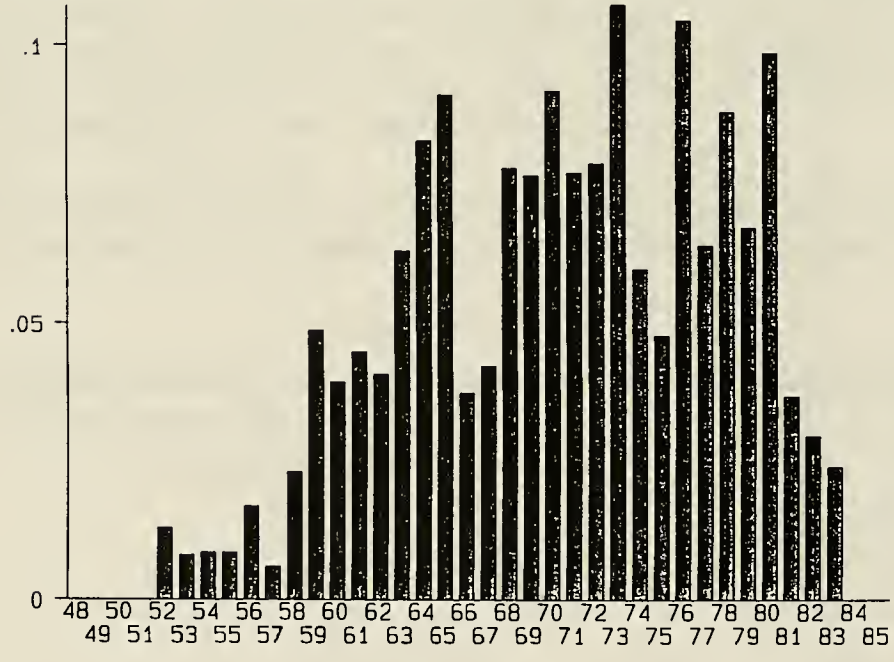


Table 1
Summary Statistics

Variable	1979 Survey			1985/86 Survey		
	all	stayers	military quitters	all	stayers	military quitters
Schooling: 9 years	0.75	0.75	0.57	0.69	0.69	0.62
10 years	0.22	0.22	0.39	0.25	0.26	0.34
12 years	0.01	0.01	0.02	0.03	0.02	0.03
13 years	0.02	0.02	0.01	0.04	0.04	0.00
Training as master craftsman	0.09	0.10	0.03	0.08	0.08	0.05
German grade: 1 (lowest)	0.02	0.02	0.03	---	---	---
2	0.11	0.11	0.15	---	---	---
3	0.47	0.47	0.39	---	---	---
4	0.34	0.34	0.39	---	---	---
5 (highest)	0.06	0.06	0.04	---	---	---
Math grade: 1 (lowest)	0.02	0.01	0.02	---	---	---
2	0.07	0.06	0.07	---	---	---
3	0.33	0.33	0.36	---	---	---
4	0.45	0.46	0.41	---	---	---
5 (highest)	0.13	0.14	0.14	---	---	---
Age	34.6	34.7	30.2	36.9	37.1	34.2
Years of experience	18.0	18.2	13.6	20.2	20.4	17.6
Severe disability	0.01	0.01	0.01	0.01	0.01	0.00
Less severe disability	0.01	0.01	0.01	0.01	0.02	0.00
Stayers	0.82	1.00	0.00	0.78	1.00	0.00
Military quitters	0.03	0.00	1.00	0.05	0.00	1.00
Still in appr. firm	0.26	0.32	0.00	0.30	0.38	0.00
Changed occupation	0.34	0.32	0.44	0.27	0.26	0.19
Changed occ. but not firm	0.04	0.04	0.03	0.03	0.03	0.01

continued

Table 1
Summary Statistics - continued

Variable	1979 Survey			1985/86 Survey		
	all	stayers	military quitters	all	stayers	military quitters
Number of employers: 1	0.25	0.31	0.10	0.34	0.37	0.32
2	0.29	0.27	0.42	0.29	0.30	0.29
3	0.23	0.22	0.27	0.18	0.17	0.21
4	0.11	0.11	0.11	0.09	0.08	0.12
5 or more	0.11	0.09	0.11	0.09	0.08	0.07
Characteristics of apprenticeship firm						
sector: manufacturing	0.31	0.33	0.22	0.33	0.36	0.30
crafts	0.43	0.42	0.43	0.41	0.39	0.42
trade	0.12	0.12	0.19	0.12	0.12	0.16
services and other	0.13	0.13	0.16	0.13	0.12	0.12
firm size: 1-9	0.33	0.30	0.33	0.31	0.28	0.32
10-99	0.35	0.35	0.41	0.39	0.39	0.45
100-999	0.19	0.20	0.17	0.20	0.21	0.17
1000+	0.13	0.14	0.10	0.10	0.12	0.06

continued

Table 1
Summary Statistics - continued

Variable	1979 Survey			1985/86 Survey		
	all	stayers	military quitters	all	stayers	military quitters
Characteristics of current job						
Hourly wage (DM)	13.08	13.12	12.88	17.14	17.30	16.84
position: semi-skld. blue col.	0.08	0.07	0.08	0.07	0.05	0.09
skilled blue collar	0.35	0.35	0.32	0.40	0.41	0.42
master craftsman	0.11	0.12	0.01	0.09	0.09	0.08
low skilled white collar	0.09	0.09	0.13	0.07	0.07	0.11
qualified white collar	0.14	0.15	0.19	0.17	0.18	0.13
highly qualified/manager	0.06	0.06	0.08	0.10	0.10	0.09
executive	0.07	0.08	0.10	0.01	0.01	0.01
self-employed	0.09	0.08	0.09	0.09	0.09	0.08
sector: manufacturing	0.43	0.44	0.38	0.43	0.44	0.34
crafts	0.21	0.21	0.16	0.24	0.24	0.28
trade	0.14	0.13	0.24	0.15	0.14	0.20
services and other	0.20	0.19	0.21	0.18	0.18	0.19
firm size: 1-9	0.20	0.19	0.21	0.23	0.22	0.26
10-99	0.33	0.33	0.41	0.35	0.36	0.43
100-999	0.26	0.27	0.19	0.26	0.27	0.19
1000+	0.21	0.22	0.19	0.16	0.16	0.12
Use apprenticeship skill: no	0.12	0.10	0.13	0.11	0.09	0.08
little	0.06	0.06	0.10	0.08	0.07	0.09
some	0.13	0.14	0.16	0.16	0.16	0.14
much	0.20	0.20	0.26	0.24	0.25	0.27
a lot	0.48	0.50	0.36	0.41	0.42	0.42
Number of observations	4,367	3,586	135	4,917	3,842	265

Note: Data are from the German Qualification and Career Survey. Samples consist of German males, age 23-59, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, employed in the private sector outside construction, and working full time.

Table 2
Basic Wage Regressions
(Dependent variable: Log average hourly earnings)

Independent variable	1979 Survey	1985/86 Survey	1984 SOEP
Years of school	0.104 (0.009)	0.076 (0.007)	0.065 (0.013)
Experience	0.114 (0.038)	0.144 (0.026)	0.113 (0.086)
Experience^2/100	-0.503 (0.316)	-0.758 (0.212)	-0.727 (0.762)
Experience^3/10,000	0.750 (1.100)	1.822 (0.698)	2.138 (2.777)
Experience^4/1,000,000	-0.029 (1.367)	-1.638 (0.806)	-2.291 (3.543)
Apprenticeship in manufacturing	0.083 (0.012)	0.100 (0.012)	---
Apprenticeship in trade	0.070 (0.018)	0.067 (0.019)	---
Apprenticeship in other sector	0.053 (0.020)	0.029 (0.019)	---
Severe disability	-0.078 (0.041)	0.000 (0.049)	0.173 (0.110)
Less severe disability	0.009 (0.046)	-0.158 (0.042)	-0.152 (0.074)
Still in apprenticeship firm	-0.032 (0.013)	-0.019 (0.012)	---
Stayer	0.022 (0.017)	0.041 (0.016)	---
Military quitter	0.045 (0.028)	0.035 (0.024)	---
Ever did military service	---	---	-0.024 (0.024)
R^2	0.106	0.132	0.130

Note: see next page

Note to Table 2: White standard errors in parentheses. Samples in the first two columns are from the Qualification and Career Survey and consist of German males, age 23-59, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, employed in the private sector outside construction, and working full time. Number of observations is 4,367 in 1979 and 4,917 in 1985/86. "Stayers" are those workers who continued in their apprenticeship firm after training; "military quitters" are those who left their training firm for military service. Sample in the last columns is from the Socioeconomic Panel and consist of German males, age 23-59, who left secondary school in 1948 or later, employed in the private sector outside construction, and working full time. Number of observations is 526. Regressions also include a constant.

Table 3
Specification Checks
(Dependent variable: Log average hourly earnings)

Specification or sample	1979 Survey		1985/86 Survey	
	stayers	military quitters	stayers	military quitters
Basic regressions	0.022 (0.017)	0.045 (0.028)	0.041 (0.016)	0.035 (0.024)
Controls for math and German grades in school	0.018 (0.017)	0.043 (0.028)	---	---
No control for still working at the apprenticeship firm	0.012 (0.016)	0.046 (0.028)	0.034 (0.015)	0.035 (0.024)
Exclude self-employed	0.013 (0.016)	0.044 (0.028)	0.022 (0.015)	0.013 (0.024)
Exclude those with 12 or more years of school	0.022 (0.017)	0.033 (0.030)	0.038 (0.016)	0.021 (0.024)
Only include those born after 1938 and include full set of age dummies	0.013 (0.020)	0.042 (0.030)	0.047 (0.018)	0.043 (0.025)
Include those who left school before 1948	0.021 (0.016)	0.008 (0.027)	0.045 (0.015)	0.049 (0.023)
Only include respondents with apprenticeship in manufacturing	-0.002 (0.036)	0.063 (0.063)	0.034 (0.030)	-0.002 (0.044)

Note: White standard errors in parentheses. Samples consist of German males, age 23-59, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, employed in the private sector outside construction, and working full time. "Stayers" are those workers who continued in their apprenticeship firm after training; "military quitters" are those who left their training firm for military service. Except where otherwise noted, additional controls are years of primary and secondary schooling, a quartic in potential experience, three dummies for the sector of the apprenticeship firm, two dummies for work related disabilities, a dummy for still working in the apprenticeship firm, and a constant.

MIT LIBRARIES
3 9080 00940 0851

Table 4
Linear Probability Models for Ability Related Outcomes

Sample	1979 Survey		1985/86 Survey	
	stayers	military quitters	stayers	military quitters
Dependent variable: Uses computer on-the-job				
Full sample	0.012 (0.011)	-0.007 (0.024)	0.036 (0.015)	0.000 (0.24)
Dependent variable: Works as shop floor supervisor ("Polier" or "Meister")				
Full sample	0.025 (0.013)	-0.043 (0.016)	0.002 (0.012)	0.001 (0.019)
Blue collar and self-employed	0.040 (0.012)	-0.024 (0.015)	0.016 (0.016)	0.003 (0.027)
Dependent variable: Trained as master craftsman				
Full sample	0.020 (0.012)	-0.034 (0.018)	-0.009 (0.012)	-0.048 (0.017)
Blue collar and self-employed	0.024 (0.017)	0.043 (0.034)	-0.005 (0.016)	-0.047 (0.023)

Note: White standard errors in parentheses. Samples consist of German males, age 23-59, who left secondary school in 1948 or later, completed private sector apprenticeship training without returning to school after the apprenticeship, employed in the private sector outside construction, and working full time. "Stayers" are those workers who continued in their apprenticeship firm after training; "military quitters" are those who left their training firm for military service. Additional controls are years of primary and secondary schooling, a quartic in potential experience, three dummies for the sector of the apprenticeship firm, two dummies for work related disabilities, a dummy for still working in the apprenticeship firm, and a constant.

Date Due

Lib-26-67